Ethnic segregation and heterogeneous preferences of homeowners for housing and neighbourhood characteristics : evidence from the Netherlands

UNU-MERIT Working Paper Series

#2013-061

Ethnic segregation and heterogeneous preferences of homeowners for housing and neighbourhood characteristics. Evidence from the Netherlands

Cheng Boon Ong and Kristof De Witte

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Maastricht Graduate School of Governance (MGSoG)
email: info-governance@maastrichtuniversity.nl | website: http://mgsog.merit.unu.edu

UNU-MERIT Working Papers

ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology, UNU-MERIT

Maastricht Graduate School of Governance MGSoG

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT and MGSoG to stimulate discussion on the issues raised.

Ethnic segregation and heterogeneous preferences of homeowners for housing and neighbourhood characteristics. Evidence from the Netherlands

Cheng Boon Ong[*†]

Kristof De Witte [α,∞]

[*] Maastricht Graduate School of Governance, Maastricht University, the Netherlands. Keizer Karelplein 19, 6211 TC Maastricht, the Netherlands. Cheng.ong@maastrichtuniversity.nl. T +31(0)433884452. F +31(0)433884499

[α] Top Institute for Evidence Based Education Research, Maastricht University, the Netherlands. Kapoenstraat 2, 6200 MD Maastricht, the Netherlands. K.dewitte@maastrichtuniversity.nl

[∞] Faculty of Economics and Business, KU Leuven. Naamsestraat 69, 3000 Leuven, Belgium. Kristof.dewitte@kuleuven.be

[†] Corresponding author.

This paper examines ethnically differentiated preferences for neighbourhood ethnic composition among homeowners in the Netherlands. Borrowing from price hedonic theory, it tests a fully nonparametric empirical model of housing choice. We exploit rich neighbourhood-level administrative data linked to the 2009 'Dutch Housing and Living Survey'. The nonparametric analysis proceeds in two steps. First, housing prices are decomposed into attribute-specific 'implicit prices'. These price hedonic estimates indicate a significant negative effect of the percentage of non-western minority residents in a neighbourhood on housing prices. For the second step and using the recovered household preference parameters, the marginal willingness to pay for an increase in non-western minority neighbours is estimated. Our model predicts an average decrease in dwelling price of €697 for every 10 per cent increase in non-western neighbours. The paper finds evidence of assimilation with some homeowners of non-western migrant background having a negative willingness to pay for living next to more co-ethnic neighbours.

Keywords: demand estimation, hedonic price, heterogeneous preference, nonparametric, generalized kernel function, ethnic segregation

JEL Classification: R20, R21, R23, R31, R32

1. Introduction

Neighbours' tastes, habits, and incomes are externalities that influence housing value (Bailey, 1959). Unlike other goods, housing is a heterogeneous good made up of utility-deriving quality attributes that include neighbourhood externalities such as public safety, crowding, pollution, composition of neighbours, and proximity to various amenities. This paper tests for ethnically differentiated preferences for neighbourhood ethnic composition. We follow the framework developed by Bajari and Kahn (2005) which relies on price hedonic modelling commonly used in real estate valuation. Housing is treated as a 'heterogeneous good' or a bundle of quality attributes such that the housing price can be decomposed into attribute-specific 'implicit prices'. Given their budget constraint, utility-maximising households are assumed to trade off various desired housing and neighbourhood attributes. The price hedonic model parameterizes housing demand and allows, in a second stage, for the estimation of households' marginal willingness to pay (MWTP) for various quality attributes including neighbourhood qualities and composition.

This paper contributes to the literature by its fully nonparametric estimation procedure. Nonparametric methods relax the parametric assumptions to accommodate the unknown functional forms that relate different housing or neighbourhood characteristics to dwelling price. Departing from Bajari and Kahn (2005), we use local linear estimation with generalized product kernels in the first-stage price hedonic estimation. We adapt the approach by mixed kernel estimations, as suggested by Racine and Li (2004; 2004), which allows for the smoothing of both categorical and continuous variables. This alleviates the problem of sparse cells due to insufficient observations.

As a second contribution, we test the microeconomic consumer choice theory model using rich Dutch housing choice data. The Netherlands makes an interesting case study for our purpose. In contrast with the segregation discourse in North American cities (Charles, 2003; Ellen, 2000; Massey and Denton, 1993; Wilson, 1987), 'ethnic segregation' in West European countries like the Netherlands revolves around a relatively homogenous ethnic majority that make up the 'nation state' and the heterogeneous immigrant minority that have arrived in the recent decades. For a country of 16.7 million, an estimated 21 per cent of Dutch population have a migrant background[1] (Centraal Bureau voor de Statistiek, 2013). In its four largest cities, this group constitute more than a third of its residents – the majority being of 'non-western' origin. The 'non-western' group is comprised of those with a Turkish, African, Latin-American and Asian background, although notably, with the exception Japanese and Indonesian. The exception is due to their higher socioeconomic position in Dutch society – implying that the 'non-western' categorization contains a socioeconomic dimension. By this definition and by historical event, 'non-western' migrants have typically assumed a low socioeconomic status when they first entered the country. This conflation of low socioeconomic status and 'non-western' background sees the persistence of ethnic segregation in housing and labour markets.

As a final contribution, the paper examines whether immigrants assimilate into the host society. Specifically, we investigate if households with a migrant background seek to

[1] The corresponding Dutch term *allochtoon* refers to someone who has at least one of her parents born abroad. A second-generation *allochtoon* is born in the Netherlands. The use of 'country of origin' as a defining characteristic in administrative data along with its academic derivatives is not without critique, see Phillips (2007) for a critical review.

purchase their homes in the same neighbourhoods as the native population households (Alba and Logan, 1992; Borjas, 2002; Logan et al., 2002). Although non-western households tend to be social housing tenants – 66 per cent in 2006 compared to 31 per cent for native Dutch households – they are increasingly becoming homeowners, up to 23 per cent as of 2006, compared to 60 per cent for native Dutch households and 43 per cent for western minority households (Ministrie VROM 2006, authors' own calculation). In a country where policymakers have actively encouraged homeownership, understanding homeowner preferences is crucial to the neighbourhood segregation debate.

The paper unfolds as follows. Section 2 provides a literature review while section 3 outlines a behaviour model to causally explain residential segregation. Section 4 discusses the estimation theory which is then followed by the data descriptive and results in Sections 5 and 6. A final section concludes the paper.

2. Literature review

We summarize the literature along three lines. The first two subsections focus on the quality attributes of housing and the influence of ethnicity on preferences. In the behavioural model outlined in Section 3, this corresponds to the first- (quality attributes) and second- stages (ethnic preferences). A third subsection provides an overview on the nonparametric estimation procedure.

2.1. Housing as a heterogeneous good

In basic consumer choice theory, the consumption good is assumed to be homogeneous. Building upon this, Houthakker (1952) introduced the notion of variety in a consumption good differentiated by its 'quality' characteristics while Lancaster (1966) redefined the idea of deriving utility from a good per se into '*supposing that it is the properties or characteristics of the goods from which utility is derived*'. The hedonic[2] price model accommodates the 'heterogeneous good', a good as a bundle of quality attributes with a price that can be decomposed into these attribute-specific implicit prices.

Appealing to 'revealed preference' theory whereby *observable* choices are said to "reveal" a consumer's *unobservable* preference set for goods given each budget constraint[3], Rosen's (1974) seminal paper introduced a two-stage estimation procedure in which the first price hedonic stage estimates the 'implicit marginal price' of each attribute by decomposing price onto the product's observed characteristics. The second stage then recovers the structural demand parameters for each attribute with the estimated marginal price and consumer characteristics[4]. Departing from Lancaster (1966), Rosen (1974) assumes that attributes cannot be unpackaged and sold separately – the indivisibility assumption – which is realistic for a durable consumption good like housing. Consequent of Rosen's (1974) theoretical work of deriving demand systems using a hedonic model, the hedonic price

[2] 'Hedonic' here appeals to the utilitarianism, as such 'hedonic price comparisons are those which recognize the potential contribution of any commodity to the welfare and happiness of its purchasers' (Court, 1939, p. 107).

[3] Revealed preference theory essentially links empirical choice data to formal consumer theory via Samuelson's weak axiom and Houthakker's strong axiom (Houthakker, 1950; for a textbook summary see Mas-Colell et al., 1995; or the seminal works of Samuelson, 1948, 1938)

[4] Rosen's second stage is often criticized for identification problems, c.f. Bartik (1987).

literature has expanded quickly to include housing markets and environmental valuation (for a recent summary of related literature see Baranzini et al., 2008).

Typically, first-stage price hedonic models involve many relevant housing and neighbourhood characteristics that influence the price such as dwelling size, number of rooms, availability of garden and garage, and so forth (see the review by Malpezzi, 2002 and references therein). Some attributes are culture, history or period-specific, which for aesthetics or (construction and maintenance) quality reasons can influence housing price (Kain and Quigley, 1970; Rubin, 1993). Unlike other goods, the spatial dimension to housing means that location matters whether it is to the central business district and other areas for access to employment, goods and services (see references in Cheshire and Sheppard, 1995; Glaeser et al., 2008). The housing good is also very much susceptible to the effects of 'local' markets (Goodman, 1978), usually defined by municipality or metropolitan statistical area. The same dwelling in terms of structural and neighbourhood characteristics could be worth differently across local housing markets due to the role of local municipality and other actors (e.g. housing associations), market size and structure (e.g. rental-homeowner stock), demography, local economy, and geographical location (for a summary of relevant literature see, De Bruyne and Van Hove, 2006).

2.2. Heterogeneous preferences and ethnic segregation

Spatial segregation between native and immigrant households has long been considered an indicator of migrant assimilation[5] into the host society (Alba and Logan, 1992; Alba et al., 1999; Logan et al., 2002; Massey, 1981; Park and Burgess, 1925). Ethnic minority households could favour living in neighbourhoods of co-ethnic concentration out of preference and taste, such as the availability of cultural-specific goods and services (Aldrich and Waldinger, 1990; Logan et al., 2002). Perception of security from having neighbours similar to one plays a role too (Farley et al., 1978). The spatial assimilation model views enclaves as transitory and predicts that, as disadvantaged minority members acculturate, they leave behind the enclaves to join neighbourhoods resided by the desirable or dominant group (Logan et al., 2002; Wilson, 1987). There is also a 'generational dynamic' with subsequent generation(s) of the minority group moving towards assimilation – for instance in their suburbanization following the dominant group's residential mobility behaviour (Alba et al., 1999; Gans, 1992; Zorlu, 2009). As pointed out by Alba and Logan (1992 p. 1315), '*homeownership, like educational and occupational advancement, residential integration or acculturation, is a potential step towards assimilation of minority group members into mainstream society*'. Since homeowners are considered as more committed than renters to their neighbourhoods (Rohe and Stewart, 1996), homeownership by households of migrant background can additionally be interpreted as their long-term commitment to the host society (Constant et al., 2009).

[5] 'Assimilation' here is used interchangeably with 'acculturation'.

Earlier sociological studies are often ecological, focused at the community or neighbourhood-level. This contrasts to microeconomic behavioural models based on consumer choice theory that attempt to causally explain residential segregation at the individual household-level (among others see Bajari and Kahn, 2008, 2005; Bayer et al., 2007; for a review, refer to Durlauf, 2004). Schelling's theoretical framework of linking neighbourhood attributes to the individual's utility function provides the basis for the latter line of residential segregation research (1972, 1971). Bajari and Kahn (2008, 2005) extend Rosen's price hedonic framework with a multi-step estimation procedure. Those in favour of the alternative method – random utility or discrete choice models (c.f. McFadden, 1978) – contend that hedonic price models assume that households can choose the level of consumption for every attribute (i.e. dense product space) when they are actually constrained by the limited housing bundles offered in the market (discussed in Bajari and Benkard, 2005; for applied examples of the discrete choice model, see Barrios García and Rodríguez Hernández, 2008; Bayer and McMillan, 2008; Wong, 2013). Nonetheless, the fact that Schelling's difference in individual preferences for neighbourhood composition can be statistically tested using housing choice data makes these models extremely appealing (Durlauf, 2004).

2.3. Nonparametric method using generalized mixed kernels

Nonparametric methods are often favoured over the more rigid parametric and semi-parametric methods due to the lack of an a priori theory as to how variables should relate to one another (Anglin and Gencay, 1996; Halvorsen and Pollakowski, 1981; Malpezzi, 2002; Parmeter et al., 2007; Triplett, 2006; Yatchew, 1998). Triplett (2006) adds that a nonparametric data-driven estimation of the hedonic function is the most appropriate method for finding each product characteristic's implicit price since '*imposing some rule for what the hedonic function "should" look like destroys part of the information that market prices convey*'. Along these lines, Bajari and Kahn (2008, 2005) utilize a flexible local linear specification of the price hedonic first stage to recover the random coefficients of housing preferences which are then modelled as a function of household demographics and household-specific shocks. In contrast, parametric specifications for the random coefficients usually assume independence and normal distribution. For a bundled good such as housing, relaxing the independence assumption with nonparametric estimation is clearly advantageous, e.g. a high valuation of dwelling size can be associated with a high valuation on the number of rooms (Bajari and Kahn, 2008).

But by not smoothing these regressors, Bajari and Kahn (2008, 2005) effectively used the 'frequency estimator' approach which splits the sample into cells based on discrete data before nonparametrically estimating the joint distributions of the continuous variables (see Racine and Li, 2004 for a more elaborate discussion). This introduces the 'curse of dimensionality' problem (Hastie and Tibshirani, 1990; Silverman, 1986), i.e. the problem of making meaningful local estimations with scarce data points in multidimensional grids as the number of dimensions (only in continuous covariates, see Racine, 2008) increases. This problem was evident in Bajari and Kahn's analysis that was restricted to seven regressors and the authors admitted that their method was subsequently not fully nonparametric (Bajari and Kahn, 2005). Semi-parametric methods could alleviate the 'curse of dimensionality' (see Hastie and Tibshirani, 1990; Lall and Lundberg, 2007). So do full nonparametric alternatives such as Racine and Li's (2004; 2004) mixed kernel estimation which allows for the smoothing of both categorical and continuous variables.

3. A theoretical model of housing choice

In the case of a heterogeneous good with heterogeneous buyers, households paying the same housing price can face different implicit prices for the various housing and neighbourhood characteristics. From the available housing choice data – i.e. the chosen bundle of housing attributes, the house price, and household characteristics – we need to infer the *unobserved* implicit prices and *unobserved* household preferences for each quality attribute. Hence, several key assumptions are required to extend consumer choice theory into estimating heterogeneous household demand or 'willingness to pay' for a specific attribute such as ethnic composition of neighbours. The next section follows the theoretical framework developed by Bajari and Benkard (2005) as applied in the housing choice context by Bajari and Kahn (2008, 2005).

Consider a model of i (1,...,*I*) households, each consuming one of the j (1,...,*J*) housing units with their respective k (1,...,*K*) attributes and deriving heterogeneous dwelling-specific utility, u_{ij}. A dwelling unit consists of its observed structural and surrounding neighbourhood characteristics, $x_{j,k}$ and unobserved attributes, ξ_j. Utility is also assumed to be time-separable in order to model housing choice as a static utility maximization problem with c as the numéraire good, and y_i as individual household income:

$$u_{ij} = \max_j u_i(x_j, \xi_j, c) \text{ subject to: } p_j + c = y_i$$

$$p_j = \mathbf{p}(x_j, \xi_j)$$

such that price, p_j, is a function of the observed and unobserved attributes, $\mathbf{p}(x, \xi)$. In equilibrium, substituting the budget constraint into the utility function allows us to find the utility maximising housing bundle, j^* for household i:

$$j^*(i) = \arg\max_j u_i\left(x_j, \xi_j, y_i - \mathbf{p}(x_j, \xi_j)\right)$$

With no adjustment costs and in a perfectly competitive market, this first-order condition of the utility function for continuous characteristics must hold:

$$\frac{\delta u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})}{\delta x_{j,k}} - \frac{\delta u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})}{\delta c}\frac{\delta \mathbf{p}(\mathbf{x}_{j^*}, \xi_{j^*})}{\delta x_{j,k}} = 0$$

$$\frac{\delta u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})/\delta x_{j,k}}{\delta u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})/\delta c} = \frac{\delta \mathbf{p}(\mathbf{x}_{j^*}, \xi_{j^*})}{\delta x_{j,k}}$$

We find that under optimality, the marginal rate of substitution between $x_{j,k}$ and c on the left-hand side is equal to the partial derivative of the hedonic function (the marginal price of $x_{j,k}$ in the market) on the right-hand side.

For identification reasons in a cross-section data setting where households are only observed once, we assume a quasi-linear functional form for the utility function with the price of the numéraire good, c set to unity (2005):

$$u_{i,j} = u(x_j, \xi_j, c) = \beta_{i,x} \ln x_j + \beta_{i,\xi} \ln \xi_j + c$$

The assumptions above let us to recover the household-specific taste parameter, $\beta_{i,k}$, for continuous housing attribute k when the household observes its utility-maximising dwelling unit, j^*.

$$\hat{\beta}_{i,k} = x_{j^*,k}\left(\frac{\partial \boldsymbol{p}(x_{j^*}, \xi_{j^*})}{\partial x_{j,k}}\right)$$

Under this specification, the household-specific taste parameter, $\beta_{i,k}$ for housing attribute k depends not only on its implicit price, but also on the consumption level of k, $x_{j^*,k}$ as part of dwelling unit j^*. Based on a threshold decision-making rule, households will only choose a dichotomous attribute when their taste parameter for it, $\beta_{i,k}$, is above its implicit price, i.e. the difference in dwelling price due solely to this attribute, $\Delta p/\Delta k$ while other attributes are held at corresponding values in x_{j^*} (Bajari and Kahn, 2005).

Heterogeneous preferences are then modelled with $\beta_{i,k}$'s interpreted as random coefficients that are a function of household characteristics d_i and orthogonal household-specific residual, η_i:

$$\beta_{i,k} = f_k(d_i) + \eta_{i,k}$$

$$\mathrm{E}(\eta_i | d_i) = 0$$

The above theoretical model will be tested according to the empirical strategy outlined in the next section.

4. Estimation Strategy

To test the theoretical model, we use insights from Bajari and Kahn's (2005) estimation procedure. This proceeds in two steps. (i) First, we estimate the hedonic price model by a fully nonparametric method with generalized smoothing kernels. This allows us to recover the household-specific taste parameters. (ii) In a second stage, using the household demographic characteristics, we estimate the marginal willingness to pay for an increase in non-western neighbours. We discuss the empirical strategy for both stages in more detail next.

4.1. First-stage: Reveal the attribute specific implicit prices

Consider a standard nonparametric model with an unknown functional form $f_j(.)$ relating the housing price, p_{j^*} to distinct housing and neighbourhood attributes, χ_k:

$$p_{j^*} = f_j(\chi_k) + \xi_{j^*}$$

The housing bundle-specific residual, ξ_{j^*} is assumed to be independent from the observed attributes, χ_k. We estimate the model using local-linear least squares estimator as introduced by Fan and Gijbels (1996).

The local density estimation requires the specification of a bandwidth, which essentially controls the bias-variance trade-off. A small bandwidth would entail a more localized estimation (the extreme being intrapolation) which reduces the bias but increases variance, while a bandwidth that is too large would oversmooth, reducing the estimator's variance while increasing its bias. While Bajari and Kahn (2005) have opted for bandwidth selection via visual inspection, we exploit the automated, data-driven fixed[6] bandwidth selection method via cross-validation proposed by Hall et al. (2004) with corrected Akaike

[6] Due to potential "spurious noise" associated with adaptive kernel estimators such as nearest-neighbour kernel estimator (Racine, 2008, p. 15) and the problem of obtaining valid bootstrapped standard errors for hypothesis testing, we employ the fixed bandwidth approach in this paper.

Information Criterion (Hurvich et al., 2002).[7] Cross-validation is computationally intensive as it involves repeated estimation of kernel density, $\hat{f}_{h,-i}(.)$ for each given i^{th} observation, X_i and bandwidth, h using all observations except the i^{th} observation. The appropriate bandwidth is chosen by minimising the integrated mean squared errors from predicting $f(X_i)$ and helps to reduce the influence of irrelevant covariates by approximating towards the bandwidth's upper bound (Hall et al., 2007; Li and Racine, 2008).

The cross-validation bandwidth selection can be sensitive to outliers and 'discretized' continuous data (Hayfield and Racine, 2008; Racine, 2008). In the application, this could be problematic for some postcode neighbourhood data provided by Statistics Netherlands that is rounded to the nearest five per cent. Hence, we omit extreme outliers[8] and treat the relatively 'discrete' continuous variables such as number of rooms or percentage of welfare benefit recipient in neighbourhood as ordered discrete variables. This reduces the number of continuous variables which would also mitigate the 'curse of dimensionality' problem discussed earlier.

4.2. Second-stage: Estimating marginal willingness to pay

In a second stage, we model the aggregated household preference distribution. This allows us to measure household-specific demand or the marginal willingness to pay (MWTP) for

[7] Although Hall et al. (2004) used least-squares cross-validation (LSCV) procedure, the AIC method is asymptotically equivalent and performs well in small samples (Li and Racine, 2004). Henderson et al. (2006) also find it to be less susceptible to the problem of under-smoothing afflicting LSCV.

[8] Using graphical methods to detect extreme outliers, we omit observations that have household income more than 10 standard deviations above the mean, or dwelling indoor size equal or more than 250m^2 or outdoor size equal or more than 300m^2. All omitted observations lie beyond the upper outer fences of the respective variable boxplots.

quality attributes. This proceeds in two steps. First, we recover household-specific preference parameters, $\hat{\beta}_{i^*,k}$. They are a function of the observed amount of attribute k consumed within the optimal dwelling unit j^*, $x_{j^*,k}$ and its estimated implicit price (2008, 2005). The hedonic price model coefficients, $\hat{\beta}_{j^*,k}$ can be interpreted as implicit prices faced by households consuming bundle j^*. Hence, for each household, the attribute-specific taste parameters are calculated as:

$$\hat{\beta}_{i^*,k} = x_{j^*,k}\left(\frac{\partial p(x_{j^*}, \xi_{j^*})}{\partial x_{j,k}}\right) = x_{j^*,k} * \hat{\beta}_{j^*,k}$$

In a second step, the MWTP is derived from the joint distribution between the taste parameters derived in the second stage, $\hat{\beta}_{i^*,k}$ and household demographics, d_i. We use the same nonparametric methods as shown in the first-stage[9] with an unknown functional form $f_k(.)$ and assuming the orthogonality of the household-specific residual, $\eta_{i,k}$. The MWTP – in the application, for a 10 per cent increase in neighbourhood proportion of non-western households – is calculated for each household using its respective taste parameter, $\hat{\beta}_{i^*,nonwest}$:

$$MWTP_i^{10\%nonwest} = \hat{\beta}_{i^*,nonwest}\left(\log\left(1.1 * \chi_{j^*,nonwest}\right) - \log\left(\chi_{j^*,nonwest}\right)\right)$$

$$= \hat{\beta}_{i^*,nonwest}(\log 1.1)$$

$$MWTP_i^{10\%nonwest} = f_k(d_i) + \eta_{i,k}$$

[9] In contrast, Bajari and Kahn (2005) use ordinary least squares regression for ease of exposition.

4.3. Hypothesis testing using bootstrapping methods

To nonparametrically test statistical significance of the predictors, we use naïve bootstrapped standard errors with the assumption of independent and identically distributed draws. Bootstrap methods are found to be superior to asymptotic methods because they have good finite-sample properties and are robust to the effect of data-driven bandwidth selection (Racine, 1997; Racine et al., 2006). In contrast to parametric linear regression where the partial derivative of the conditional mean with respect to a continuous predictor is assumed to be constant over the entire domain, nonparametric regression allows for the predictor's vector of partial derivatives to vary over its domain. Hence, the corresponding null hypothesis tests if these partial derivatives are equal to zero for the entire domain (Racine, 1997). Similarly, the null hypothesis for categorical regressors would be that the conditional mean with respect to the categorical variable z and a vector of other regressors $\boldsymbol{x}$ are equal to the conditional mean with respect to only $\boldsymbol{x}$ for 'almost everywhere' (Racine et al., 2006). Analogous to the standard t-tests in parametric regression, we refer to Racine (1997) and Racine et al. (2006) for the technical details with regards to hypothesis testing of continuous and categorical variables respectively.

5. Data

We use the 2009 national Dutch Housing Survey (*Woononderzoek Nederland*, WoON)[10] dataset with information on household characteristics along with housing attributes, preferences and mobility for over 70,000 respondents (Centraal Bureau voor de Statistiek,

[10] This "2009_r_1.4" version (Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieubeheer (VROM), 2009) was provided via the Data Archiving and Networked Services of the Royal Netherlands Academy of Arts and Sciences.

2010). While we have conducted our analyses for the four largest cities in the Netherlands – Amsterdam, Rotterdam, The Hague, and Utrecht – within the confines of this paper, we limit our discussion of the results to the municipality of The Hague which was oversampled that year with a total of 5,382 respondents, only 1,963 of which are homeowners[11]. This much larger sample, by a factor of three compared to the other cities, produce robust estimates using the data-driven nonparametric method. Since we assume that households take neighbourhood characteristics as given while positive transaction costs of moving could hinder dissatisfied households from consuming their utility-maximising housing bundles in the short run, we limit our sample to homeowners who have lived in their dwellings for less than 10 years[12]. Excluding missing values, outliers, and those who have lived in their present dwellings for 10 years and beyond, the final analyses involve 1,145 homeowners in The Hague.

The dependent variable for the price hedonic estimation is the *self-reported* housing price. Malpezzi (2002) posits that 'while the variances of owner assessments are high, biases are modest', hence owner-assessed housing price can be reliable given sufficient data. The alternative would be to use the imputed dwelling price, *WOZ*-value (*waardering onroerende zaken waarde*) used by the local municipality to assess property tax per dwelling but we risk systematic bias from data imputation.

[11] Results for all four cities are available upon request.

[12] Similar studies (Bajari and Kahn, 2005; DiPasquale and Kahn, 1999) have limit their sample to households who were living in the same dwelling for the last five years. According to our dataset, less than 2 percent of homeowners were dissatisfied with their current dwelling so we assume that transaction costs of moving to be negligible and that there were no major shocks to neighbourhood composition during this period – using administrative neighbourhood data, we found a median change in the number of non-western minority in a neighbourhood (as a proportion of total number of residents) to be around 5 percent between 1998 and 2008 (mean = 0.075 standard deviation = 0.076).

With the four-digit postcode identifier, household-level information is then complemented with neighbourhood-level administrative data. The common operational definition of a 'neighbourhood' used in Dutch housing studies is based on the four-digit unique postcodes with an average of 4,000 inhabitants[13]. Such administrative data is reliable as every resident in the country is obliged to register her most current address with the local municipality of residence. Neighbourhood variables include the proportion of households from different ethnic background, the degree of urbanization, average household size, and the percentage of households with children.

For the first-stage estimation, we have included relevant covariates according to the international and Dutch-specific housing price hedonic literature. This includes variables found in a standard real estate vacancy announcement: dwelling type (detached, semi-detached, corner, or terrace house and apartment), indoor and outdoor size, number of rooms, and availability of garden, car park, balcony, and central heating. Additionally, we control for the Dutch-specific single-storey access of primary dwelling chambers and the dwelling's construction period (pre-war or 1945, between 1945 and 1959, 1960 and 1969, 1970 and 1979, 1980 and 1989, 1990 and 1999, and 2000 and beyond). Besides our main neighbourhood variable-of-interest, share of non-western residents in a neighbourhood, we also include socioeconomic neighbourhood variables – average income per income-earner and the share of welfare benefit recipient – to control for household preference for

[13] The neighbourhood data is provided at the administratively defined 'neighbourhood' or '*buurt*'-level which is smaller than the statistical definition of neighbourhood based on the four-digit postcode. For our analysis, these neighbourhood variables have been aggregated to the postcode-level.

neighbourhood characteristics beyond ethnicity. Although the dataset lacks information on commuting time, focusing on one city offsets this shortcoming. Moreover, we control for the type of neighbourhood ('urban-central', 'urban-outer-central', 'urban-green') that measures degree of urbanization and proxies for 'locality' in the sense of distance to the city centre. All continuous variables are natural log-transformed as prescribed by Bajari and Benkard's utility function specification (2005).

As for the second-stage estimation of marginal willingness to pay for an increase in non-western neighbours, we control for life-cycle demographic effect on housing demand using proxy variables: the head of household's age, household type (single, couple without children, couple with children, single parent with children, other) and size. To avoid omitted variable bias, we also account for the household's highest attained level of education and (standardized) disposable income – both variables most likely correlated with ethnicity and have an effect on household preference for neighbourhood composition. Our categorization of 'ethnicity' in a broad sense includes generational status and is divided into: native Dutch, first-generation non-western, first-generation western, second-generation non-western and second-generation western.

The descriptive statistics of housing attributes by ethnicity are provided in *Table 1*, while household characteristics' descriptive statistics are provided in *Table 2*.

Table 1*: Housing characteristics by ethnicity (in column percentages)*

	Native Dutch	Non-western	Western	Total
Mean self-reported housing price (€)	**233914**	**203359**	**260910**	230340
Mean indoor dwelling size (m^2)	**111**	**110**	**124**	113
Mean outdoor dwelling size (m^2)	**70**	**48**	**75**	65
Mean number of rooms	**4**	**4**	**4**	4
Type of dwelling				
Detached house	**1.5**	**2.4**	**3.2**	2.0
Semi-detached house	**4.2**	**2.8**	**10.4**	4.8
Corner house	**6.7**	**6.7**	**5.8**	6.6
Terrace house	**24.3**	**29.6**	**24.1**	25.6
Apartment	**63.4**	**58.5**	**56.5**	61.1
Construction period				
pre-1945	**37.6**	**35.7**	**37.1**	37.0
1945 to 1959	**8.1**	**5.9**	**8.3**	7.6
1960 to 1969	**14.3**	**4.8**	**13.3**	11.7
1970 to 1979	**4.2**	**3.3**	**1.8**	3.6
1980 to 1989	**2.2**	**2.0**	**3.6**	2.4
1990 to 1999	**14.1**	**21.5**	**17.3**	16.4
2000 and beyond	**19.5**	**27.0**	**18.7**	21.3
With shared/attached garden	**52.5**	**52.0**	**63.7**	54.1
No shared/attached garden	**47.5**	**48.0**	**36.3**	45.9
With balcony	64.5	62.8	59.0	63.2
No balcony	35.5	37.2	41.0	36.8
With car park/garage	**16.0**	**15.0**	**21.9**	16.6
No car park/garage	**84.0**	**85.0**	**78.1**	83.4
With central heating	**69.6**	**83.9**	**69.8**	73.2
No central heating	**30.4**	**16.1**	**30.2**	26.8
Single-storey access to main rooms	**41.6**	**46.5**	**49.3**	44.0
No single-storey access	**58.4**	**53.5**	**50.7**	56.0
Neighbourhood type				
Urban-central	**2.8**	**2.0**	**2.5**	2.6
Urban-outer central	**90.5**	**93.9**	**88.1**	91.0
Urban-green	**6.7**	**4.1**	**9.0**	6.4
Mean share of non-western residents	**30.4**	**50.1**	**30.9**	35.4
Mean income per earner (€)	**30299**	**25346**	**31518**	29241
Mean share of welfare-benefit recipients	**15.5**	**18.9**	**15.4**	16.3
Number of observations	1095	460	278	1833

Source: Dutch Housing and Living Survey (WoON) 2009 for The Hague, authors own calculation. Bolded values represent statistically significant tests of independence at the five per cent level based on the chi-square statistic for categorical variables and the one-way analysis of variance (ANOVA) Bartlett statistic for continuous variables.

Table 2: Household characteristics by ethnicity of household head (column percentages)

	Native Dutch	*Non-western*	*Western*	**Total**
Household type				
Single household	**34.3**	**14.8**	**23.0**	27.7
Couple with no children	**32.1**	**17.6**	**30.2**	28.2
Couple with children	**27.4**	**53.3**	**41.7**	36.1
Single parent	**3.9**	**10.4**	**4.3**	5.6
Non-family	**2.4**	**3.9**	**0.7**	2.5
Level of education				
Primary	**1.28**	**10.7**	**1.4**	3.66
Lower secondary	**15.6**	**16.5**	**9.0**	14.84
Higher secondary	**31.1**	**36.3**	**29.9**	32.24
Tertiary	**52.0**	**36.3**	**59.7**	49.21
Mean household size	**2.2**	**3.1**	**2.6**	2.5
Mean household income (€)	**37395**	**35980**	**45018**	38202
Mean Age of household head	**47**	**41**	**49**	46
Number of observations	1095	460	278	1833

Source: Dutch Housing and Living Survey (WoON) 2009 for The Hague, authors own calculation. Bolded values represent statistically significant tests of independence at the five per cent level based on the chi-square statistic for categorical variables and the one-way analysis of variance (ANOVA) Bartlett statistic for continuous variables.

Non-western minority homeowners in The Hague disproportionately live in newer housing units built after 1990. They seem to value outdoor dwelling size a lot less than native Dutch and western minority homeowners. On average, both native Dutch and western minority groups live in neighbourhoods where a third of its residents are of non-western background while non-western homeowners live in neighbourhoods where half are (in broad terms) 'co-ethnic' neighbours. Given these noticeable differences at the exploratory stage, we seek to see if they persist when control variables are accounted for with the results of our explanatory model presented in the next section.

6. Results

6.1. First stage: Reveal the attribute specific implicit prices

In the conventional, linear-additive parametric model, an explanatory variable's coefficient is interpreted as its average effect on the dependent variable while holding the other covariates constant. This is, however, more complicated in the case of the full nonparametric model as the effect of an explanatory variable varies depending on the values of the other covariates. Without holding other predictors constant or at specific values, it is more useful to view the effect of an explanatory variable as a distribution rather than a point-estimate and report the summary measures for each of them. The nonparametric results for the first-stage housing price hedonic, i.e. the implicit prices, are provided in *Table 3*. As much as 92 per cent of the variance in housing price is explained by the predictors. All the independent variables are found to be statistically significant at the five per cent level using independent and identically distributed bootstrapped standard errors.

Table 3: First-stage: Price hedonic results with housing price as the dependent variable

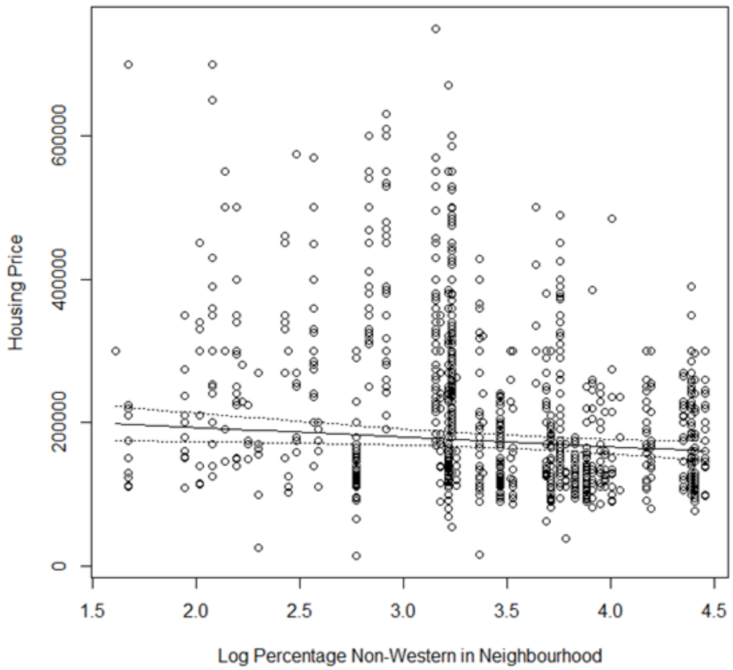
Housing price	*Mean*	*Q1*	*Median*	*Q3*	p-value
(log) % non-western minority	**-16848**	**-27146**	**-13683**	**-489**	**0.0276**
Construction period	963	-902	0	2682	0.0000
Number of rooms	1096	-886	253	2320	0.0000
(log) Indoor dwelling size	104487	65196	88575	131086	0.0000
(log) Outdoor dwelling size	13993	6370	9510	16159	0.0050
Single-storey	766	0	0	0	0.0000
Garden	235	0	0	308	0.0000
Balcony	-2031	-4913	0	0	0.0000
Car park	5858	0	0	0	0.0000
Dwelling type	-7037	-12181	-3334	452	0.0050
Central heating	308	-545	0	827	0.0000
% welfare benefit recipient	-472	-1550	-400	733	0.0000
(log) Mean income per earner	46490	-2751	32289	94357	0.0000
Living environment type	-527	-2027	-383	706	0.0000

Note: *Q1* and *Q3* refer to the first and third quartile respectively. Bootstrapped p-values are from 399 replications using independent and identically distributed draws (i.i.d.). The R^2statistic is 0.92 while sample size equals 1145 observations.

Our model predicts an average decrease in dwelling price of €697 for every 10 per cent increase in non-western neighbours[14]. A more specific graphical presentation is given in *Figure 1,* which displays the scatter plot and partial regression plot of housing price on log percentage of non-western neighbours (with 95 per cent confidence interval bands) when other housing and neighbourhood variables are held at their median or mode values. In the case of our homeowner sample for The Hague, the mode dwelling is an old, pre-1945 apartment with central heating measuring 100m^2 in indoor space and 20m^2 in outdoor space, a garden and balcony but without a car park, four rooms that do not share single-storey access, and is located in a 'urban outer-central' neighbourhood with an average income (per income earner) of €26,500 where 29 per cent of its residents are of non-western minority background and 17.5 per cent receive welfare benefits. The linear decreasing partial regression function in *Figure 1* corresponds to *Table 3* – that is, the average effect of non-western neighbours on housing price is negative when holding other covariates at their median and mode values. In addition, we note that the linear relationship here is imposed by the natural logarithmic transformation of the regressor based on our theoretical model.

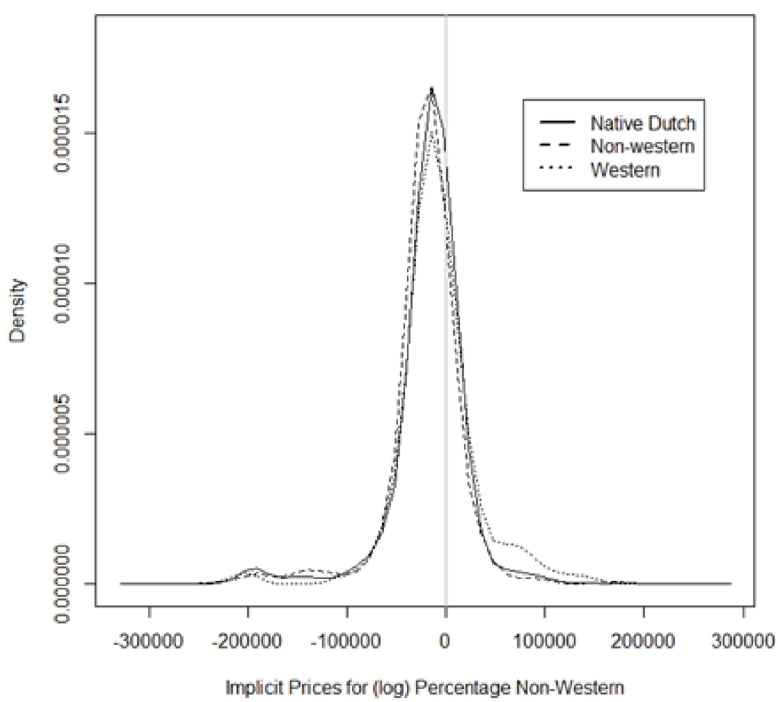
[14] The marginal willingness to pay for a 10 percent increase in non-western neighbours is calculated as $\hat{\beta}_{i^*,nonwest} * (\log 1.1)$. $\hat{\beta}_{i^*,nonwest}$ in this case is estimated to be -16848 (see *Table 3*).

Figure 1: Implicit price of (log) percentage non-western neighbours



Note: Pointwise 95 per cent confidence intervals are estimated based on the percentile (0.025, 0.975) bootstrap distribution from 100 replications.

Figure 2: Kernel distribution of implicit prices for (log) percentage of non-western minority by ethnicity in The Hague



For the different ethnicity groups, we present in *Figure 2* the kernel distribution of the coefficients which one can collectively interpret as the implicit price for (log) percentage of non-western residents in the neighbourhood.[15] It is a descriptive way of checking for potential heterogeneity in the demand for non-western neighbours (i.e. the larger the difference between the distributions, the more heterogeneity exists between the groups). Native Dutch and western minority households observe similar density functions (except for a fatter right tail distribution for western minority), while slightly more non-western households have a negative demand for non-western neighbours. As expected from the aggregated implicit price for non-western neighbours, we see that the majority of households have a negative demand for non-western neighbours.

Concerning the other control variables in *Table 3*, we observe some intuitive estimation results. Homeowners have a higher willingness to pay if the house has more rooms, is larger (both inside and outside), and has single-story access to its primary chambers, a garden, car park and central heating. Finally, the higher the mean income per earner in a neighbourhood, the higher the mean housing price.

6.2. Second stage: Estimating marginal willingness to pay

The nonparametric results for the second-stage estimation for households' marginal willingness to pay (MWTP) for a 10 per cent increase in non-western neighbours are provided in *Table 4* below. Similar to the previous table of results, *Table 4* reports the

[15] As comparison, kernel distribution of implicit prices for percentage of non-western minority by level of education is produced in *Figure B* in the *Appendix.*

summary measures (e.g. mean and median) for the effect of household characteristics. Due to the log specification the dependent variable can be interpreted as a constant percentage change of a proportion, e.g. from 8 per cent to 8.8 per cent or from 80 per cent to 88 per cent.

The results indicate that there is a statistically significant effect of household ethnicity and highest attained level of education on their preference for non-western neighbours. The mean effect of ethnicity across the various subgroups – first- and second-generation, non-western and western[16] – compared to the reference native Dutch category is -€4126. In other words, non-native Dutch homeowners in The Hague have a negative marginal willingness to pay of €4126 for a 10 per cent increase in non-western neighbours. We also observe that at least a quartile (but less than half since the median is zero) of non-western and western minority homeowners in The Hague observe negative preferences for more non-western neighbours. The median effect of education attainment is negative on their willingness to pay for more non-western neighbours. Age of household head has a clear positive effect on household preference for non-western neighbours[17] but we do not find statistical significant results for household type and household income. The insignificant finding for homeowners in The Hague is in contrast with the results of Bolt and van Kempen (2010) who have found that native Dutch couple households with children and higher income households are far more likely to move into 'non-concentrated'

[16] The order of the ethnic subgroups besides the first (for the reference category) does not matter since 'ethnicity' here is entered as an unordered categorical variable.
[17] The positive age effect is expected since there are disproportionately more native Dutch homeowners who are both disproportionately older and more likely (compared to other ethnic groups) to be residing in neighbourhoods with more non-western households.

neighbourhoods with less than 40 per cent non-western residents (instead of 'concentrated' neighbourhoods).

Table 4: Second stage: Household MWTP for 10% increase in non-western neighbours

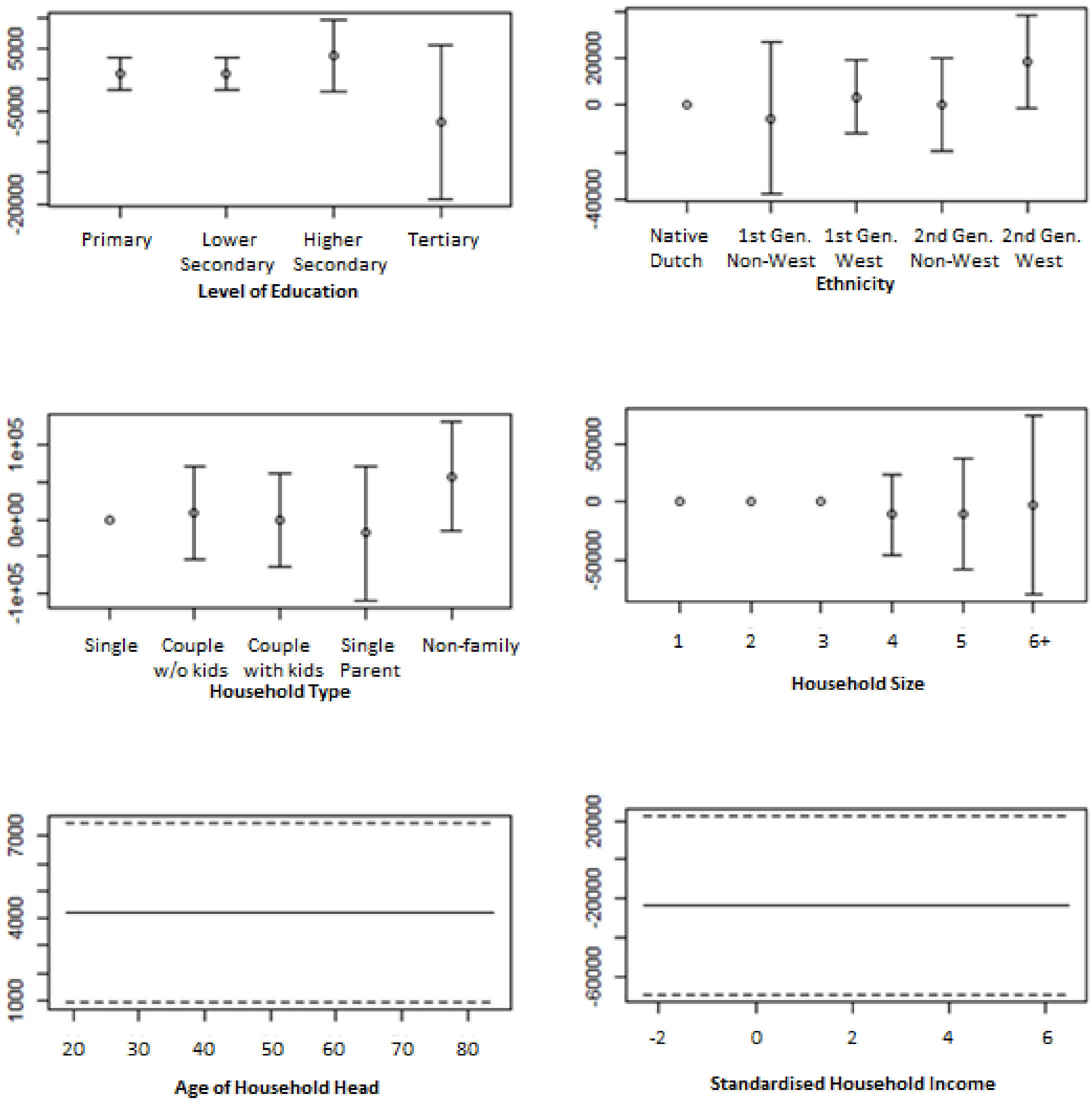
MWTP for 10% increase in non-western	*Mean*	*Q1*	*Median*	*Q3*	p-value
Ethnicity	-4126	-4790	0	0	0.0025
Level of education	-1272	-4384	-565	2970	0.0451
Household size	-1308	0	0	0	0.0326
Household type	1150	-8109	0	13701	0.3133
Household income	-3960	-22035	-7569	19191	0.3133
Age of household head	2130	971	1229	2509	0.0000

Note: *Q1* and *Q3* refer to the first and third quartile respectively. Bootstrapped p-values are from 399 replications using independent and identically distributed draws (i.i.d.). The R^2statistic is 0.18 while sample size equals 1145 excluding missing values and outliers.

To examine the conditional effects, we use the partial gradient plots in *Figure 3* holding other covariates are at their median or mode values. The mode household in our sample is a native Dutch couple with children[18] with at least one parent being tertiary-level educated, its head of household aged 38, and disposable household income slightly below (0.3 standard deviation) the national household mean income. Homeowners of second-generation western background exhibit positive preference for more non-western neighbours compared to native Dutch households. Non-western minority households themselves, however, do not exhibit this preference for their (in broad terms) 'co-ethnics' when other predictors are held at their median or mode values. The first-generation non-western homeowners appear to have a slightly negative mean preference (but with large overlap in error bars) compared to native Dutch and other households.

[18] We note that the median household size is 2 which is inconsistent with our mode household type of a couple with children, hence both variables in *Tables 4* and *5* in the partial gradient plots should be interpreted with care. Nonetheless, the bootstrapped standard errors are still valid for hypothesis testing.

Figure 3: Partial gradient plots of willingness to pay for more non-western neighbours



Note: Partial gradient plots while other covariates are held at their median or mode values. Pointwise 95 per cent confidence intervals are estimated based on the percentile (0.025, 0.975) bootstrap distribution from 100 replications.

So far we have collectively tested the ethnicity effect across the various subgroups using bootstrapped standard errors. To specifically test the effect of 'non-western' background on homeowner preference for non-western neighbours, we estimate an additional second-stage model using separate ‘non-western’ and ‘western’ ethnic variables. The summary measures of the MWTP by household characteristics are reported in *Table 5*. Compared to native Dutch and western minority homeowners, non-western homeowners in The Hague have a negative marginal willingness to pay of €6110 for a 10 per cent increase in non-western neighbours. At least a quartile (but less than half since the median is zero) of non-western minority homeowners in The Hague observe negative preferences for more non-western neighbours. Hence, we argue that our results provide clear evidence of assimilation by *some* non-western households to purchase their homes in neighbourhoods with fewer co-ethnics, even after controlling for various housing, neighbourhood and household characteristics within our two-stage nonparametric framework. Previous findings in the Netherlands combining both the homeowner and renter samples have found that non-western minority households prefer to live with ‘co-ethnic’ neighbours (Bolt et al., 2008; van Ham and Feijten, 2008). While they have controlled for homeownership, they have not accounted for the interaction between homeownership and ethnicity which could partially explain the contradicting conclusions.

Table 5: Second stage: Household MWTP for 10% increase in non-western neighbours

MWTP for 10% increase in non-western	*Mean*	*Q1*	*Median*	*Q3*	p-value
Non-western	-6110	-2942	0	0	0.0226
Western	2058	0	0	0	0.0426
Level of education	-1390	-5598	-847	3615	0.0727
Household size	-1580	0	0	0	0.0401
Household type	1078	-8882	0	13376	0.2782
Household income	-4183	-20281	-8928	17029	0.6065
Age of household head	1934	851	1232	1986	0.0050

Note: *Q1* and *Q3* refer to the first and third quartile respectively. Bootstrapped p-values are from 399 replications using independent and identically distributed draws (i.i.d.). The R^2 statistic is 0.17 while sample size equals 1145 excluding missing values and outliers.

7. Conclusion

Previous research observed that households of migrant background have a preference or higher tolerance of living with their co-ethnics or other migrant groups, compared to the native or dominant group. This paper contributes to this literature by taking into account the bundling nature of housing and neighbourhood attributes within a housing market, i.e. a household does not choose neighbourhood composition per se but rather trades it off with many other characteristics of different importance to the household. Using a flexible nonparametric framework makes our price hedonic model less susceptible to misspecification issues of parametric models. We apply a two-stage nonparametric framework to rich housing choice data for the Dutch administrative capital, The Hague.

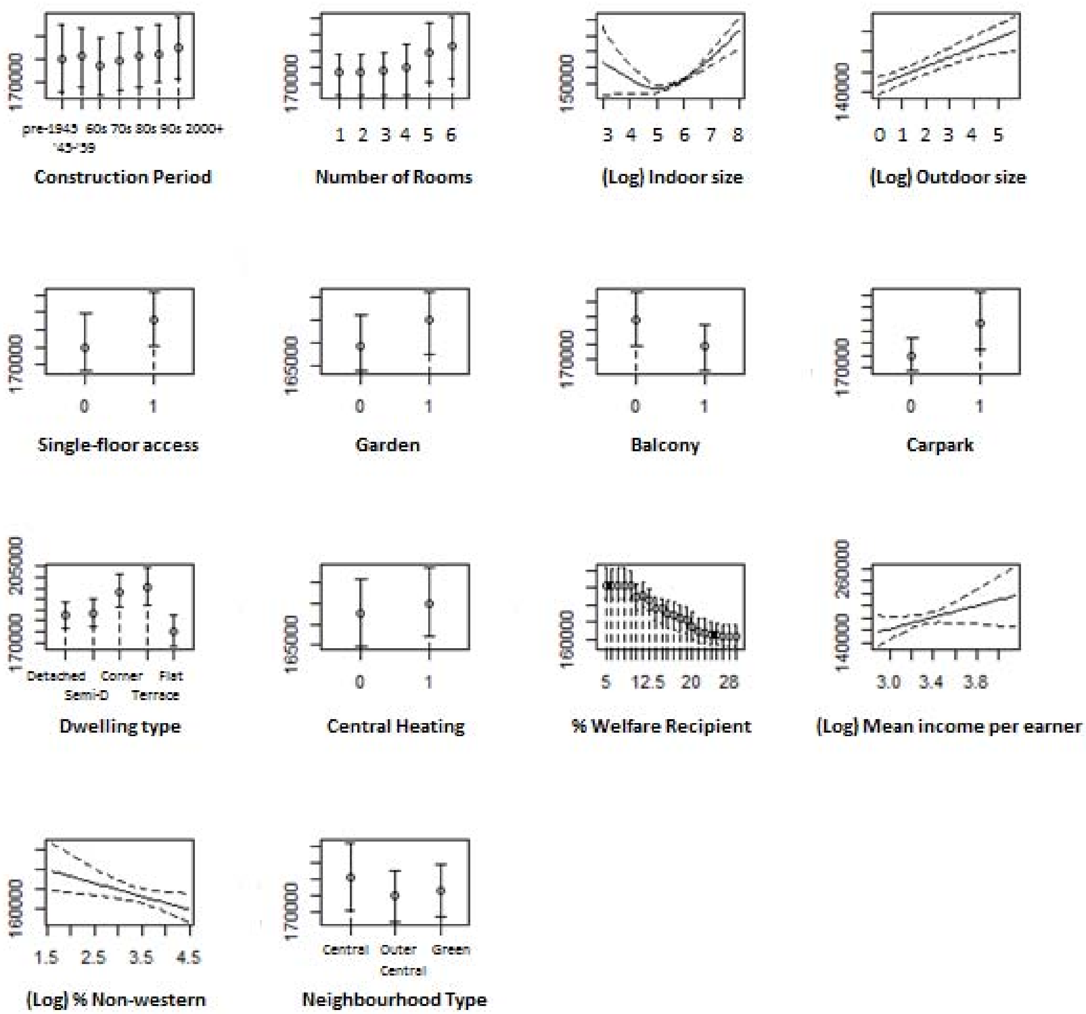
Our model predicts an average decrease in dwelling price of €697 for every 10 per cent increase in non-western neighbours when other housing and neighbourhood variables are held at their median or mode values. More importantly, we find statistically significant difference in household preference between native Dutch, non-western minority, and western minority homeowners. While the median non-western homeowner is relatively indifferent towards neighbourhood ethnic composition, the results suggest that at least some prefer *not* to live with more non-western households. This evidence of assimilation by non-western minority households who could afford homeownership is in line with the many fiscal and urban renewal policies of the Dutch government to encourage homeownership and expand the homeowners housing sector at the expense of the social rented sector. For ethnically integrated neighbourhoods, this finding supports the present Dutch policies of increasing the supply of home-owned dwellings and the rate of homeownership in the country. Further research on self-selection into homeownership and how to encourage homeownership among the relatively disadvantaged non-western minority groups could assist policymakers into reducing barriers of socioeconomic mobility and spatial assimilation.

Acknowledgements

We are indebted to Jan Rouwendal for inspiring this paper and his helpful comments. We would also like to thank Henriëtte Maassen van den Brink, Aslan Zorlu, Remi de Cock, and Mindel van de Laar for guidance and support. This study would not have been possible without the data from Statistics Netherlands (CBS) and the (then) Ministry of Housing, Spatial Planning and the Environment (VROM) of the Netherlands. The usual caveat applies.

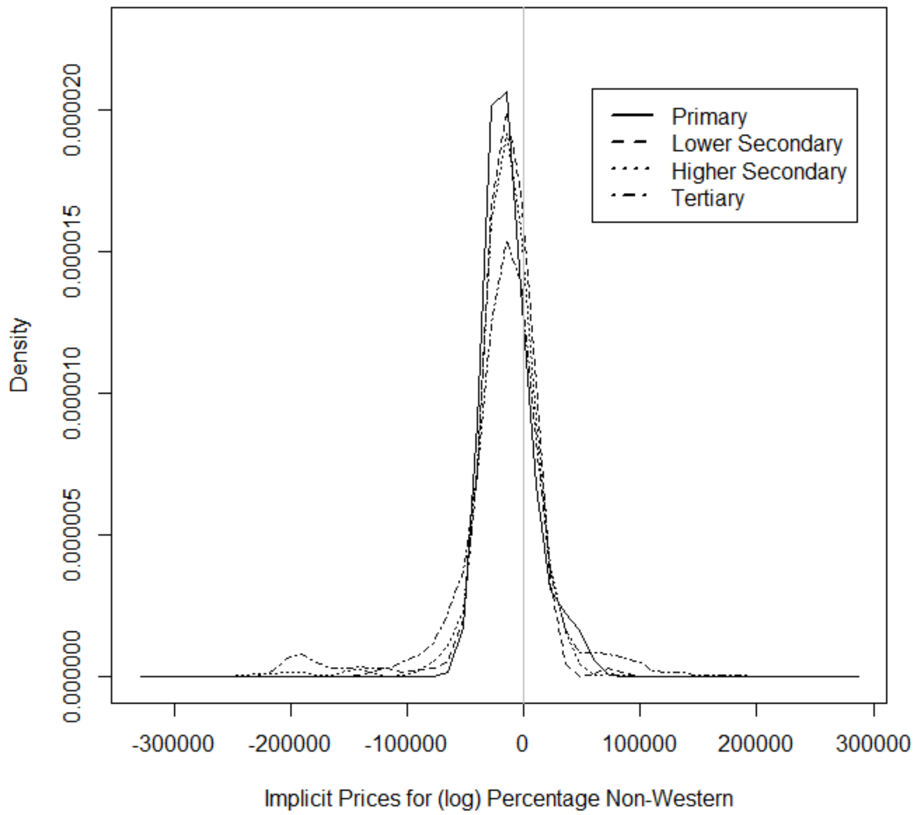
APPENDIX

Figure A: Partial regression plots for housing and neighbourhood attributes



Note: Partial regression plots while other covariates are held at their median or mode values. Pointwise 95 per cent confidence intervals are estimated based on the percentile (0.025, 0.975) bootstrap distribution from 100 replications.

Figure B: Kernel distribution of implicit prices for (log) percentage of non-western minority by highest attained level of education



BIBLIOGRAPHY

Alba, R.D., Logan, J.R., 1992. Assimilation and stratification in the homeownership patterns of racial and ethnic groups. International Migration Review 26, 1314–1341.

Alba, R.D., Logan, J.R., Stults, B.J., Marzan, G., Zhang, W., 1999. Immigrant groups in the suburbs: A reexamination of suburbanization and spatial assimilation. American sociological review 446–460.

Aldrich, H.E., Waldinger, R., 1990. Ethnicity and entrepreneurship. Annual review of sociology 16, 111–135.

Anglin, P.M., Gencay, R., 1996. Semiparametric estimation of a hedonic price function. Journal of Applied Econometrics 11, 633–648.

Bailey, M.J., 1959. Note on the economics of residential zoning and urban renewal. Land Economics 35, 288–292.

Bajari, P., Benkard, C.L., 2005. Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. Journal of Political Economy 113, 1239–1276.

Bajari, P., Kahn, M.E., 2005. Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. Journal of Business and Economic Statistics 23, 19–33.

Bajari, P., Kahn, M.E., 2008. Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl, in: Baranzini, A., Ramirez, J.V., Schaerer, C., Thalmann, P. (Eds.), Hedonic Methods in Housing Markets. Pricing Environmental Amenities and Segregation. Springer, New York, pp. 128–155.

Baranzini, A., Ramirez, J.V., Schaerer, C., Thalmann, P., 2008. Hedonic Methods in Housing Markets. Pricing Environmental Amenities and Segregation. Springer, New York.

Barrios García, J.A., Rodríguez Hernández, J.E., 2008. Housing demand in Spain according to dwelling type: Microeconometric evidence. Regional Science and Urban Economics 38, 363–377.

Bartik, T.J., 1987. The estimation of demand parameters in hedonic price models. The Journal of Political Economy 95, 81–88.

Bayer, P., Ferreira, F., McMillan, R., 2007. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. Journal of Political Economy 115, 588–638.

Bayer, P., McMillan, R., 2008. Distinguishing Racial Preferences in the Housing Market: Theory and Evidence, in: Baranzini, A., Ramirez, J.V., Schaerer, C., Thalmann, P. (Eds.), Hedonic Methods in Housing Markets. Pricing Environmental Amenities and Segregation. Springer, New York, pp. 225–245.

Bolt, G., van Kempen, R., 2010. Ethnic Segregation and Residential Mobility: Relocations of Minority Ethnic Groups in the Netherlands. Journal of Ethnic and Migration Studies 36, 333–354.

Bolt, G., van Kempen, R., van Ham, M., 2008. Minority Ethnic Groups in the Dutch Housing Market: Spatial Segregation, Relocation Dynamics and Housing Policy. Urban Studies 45, 1359.

Borjas, G.J., 2002. Homeownership in the immigrant population. Journal of Urban Economics 52, 448–476.
Centraal Bureau voor de Statistiek, 2010. Onderzoeksverantwoording WOON 2009, Dataverzameling Woningmarktmodule. Centraal Bureau voor de Statistiek (CBS), Den Haag/Heerlen.
Centraal Bureau voor de Statistiek, 2013. Statline: Bevolking Kerncijfers [WWW Document]. URL http://statline.cbs.nl/ (accessed 3.25.13).
Charles, C.Z., 2003. The Dynamics of Racial Residential Segregation. Annual review of sociology 29, 167–207.
Cheshire, P., Sheppard, S., 1995. On the price of land and the value of amenities. Economica 247–267.
Constant, A.F., Roberts, R., Zimmermann, K.F., 2009. Ethnic identity and immigrant homeownership. Urban Studies 46, 1879–1898.
Court, A.T., 1939. Hedonic price indexes with automotive examples, in: The Dynamics of Automobile Demand. General Motors Company, New York, pp. 99–117.
De Bruyne, K., Van Hove, J., 2006. Explaining the spatial variation in housing prices: an economic geography approach. Applied Economics 45, 1673–1689.
DiPasquale, D., Kahn, M.E., 1999. Measuring neighborhood investments: An examination of community choice. Real Estate Economics 27, 389–424.
Durlauf, S.N., 2004. Neighborhood effects, in: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics. Elsevier, Amsterdam, pp. 2173–2242.
Ellen, I.G., 2000. Sharing America's neighborhoods: The prospects for stable racial integration. Harvard University Press, Cambridge, MA.
Fan, J., Gijbels, I., 1996. Local polynomial modelling and its applications, Vol. 66 of Monographs on Statistics and Applied Probability. London: Chapman Hall.
Farley, R., Schuman, H., Bianchi, S., Colasanto, D., Hatchett, S., 1978. Chocolate city, vanilla suburbs: will the trend toward racially separate communities continue? Social Science Research 7, 319–344.
Gans, H.J., 1992. Second-generation decline: scenarios for the economic and ethnic futures of the post-1965 American immigrants. Ethnic and racial studies 15, 173–192.
Glaeser, E.L., Kahn, M.E., Rappaport, J., 2008. Why do the poor live in cities The role of public transportation. Journal of Urban Economics 63, 1–24.
Goodman, A.C., 1978. Hedonic prices, price indices and housing markets. Journal of Urban Economics 5, 471–484.
Hall, P., Li, Q., Racine, J.S., 2007. Nonparametric estimation of regression functions in the presence of irrelevant regressors. The Review of Economics and Statistics 89, 784–789.
Hall, P., Racine, J.S., Li, Q., 2004. Cross-validation and the estimation of conditional probability densities. Journal of the American Statistical Association 99, 1015–1026.
Halvorsen, R., Pollakowski, H.O., 1981. Choice of functional form for hedonic price equations. Journal of Urban Economics 10, 37–49.
Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Monographs on Statistics and Applied Probability 43.
Hayfield, T., Racine, J.S., 2008. Nonparametric Econometrics: The np Package. Journal of Statistical Software 27, 1–32.

Henderson, D.J., Olbrecht, A., Polachek, S.W., 2006. Do former college athletes earn more at work? A nonparametric assessment. Journal of Human Resources 41, 558–577.
Houthakker, H.S., 1950. Revealed preference and the utility function. Economica 17, 159–174.
Houthakker, H.S., 1952. Compensated changes in quantities and qualities consumed. The Review of Economic Studies 19, 155–164.
Hurvich, C.M., Simonoff, J.S., Tsai, C.-L., 2002. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 60, 271–293.
Kain, J.F., Quigley, J.M., 1970. Measuring the value of housing quality. Journal of the American Statistical Association 65, 532–548.
Lall, S.V., Lundberg, M., 2007. What are public services worth, and to whom? Non-parametric estimation of capitalization in Pune. Journal of Housing Economics.
Lancaster, K.J., 1966. A new approach to consumer theory. The Journal of Political Economy 74, 132.
Li, Q., Racine, J.S., 2004. Cross-validated local linear nonparametric regression. Statistica Sinica 14, 485–512.
Li, Q., Racine, J.S., 2008. Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. Journal of Business & Economic Statistics 26, 423–434.
Logan, J.R., Zhang, W., Alba, R.D., 2002. Immigrant enclaves and ethnic communities in New York and Los Angeles. American Sociological Review 299–322.
Malpezzi, S., 2002. Hedonic Pricing Models: A selective and applied review, in: Gibb, K., O'Sullivan, A. (Eds.), Housing Economics and Public Policy. Blackwell, Oxford.
Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, New York/Oxford.
Massey, D.S., 1981. Dimensions of the New Immigration to the United States and the Prospects for Assimilation. Annual Review of Sociology 7, 57–85.
Massey, D.S., Denton, N.A., 1993. American Apartheid: Segregation and the Making of the Underclass. Harvard University Press, Cambridge, MA.
McFadden, D., 1978. Modelling the Choice of Residential Location, in: Karlqvist, A., Lundqvist, L., Snickers, F., Weibull, J.W. (Eds.), Spatial Interaction Theory and Planning Models. North-Holland, New York.
Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieubeheer (VROM), 2006. WoonOnderzoek Nederland (WoON) [WWW Document]. URL http://easy.dans.knaw.nl/
Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieubeheer (VROM), 2009. WoonOnderzoek Nederland (WoON) [WWW Document]. URL http://easy.dans.knaw.nl/
Park, R.E., Burgess, E.W., 1925. The City. University of Chicago Press, Chicago, IL.
Parmeter, C.F., Henderson, D.J., Kumbhakar, S.C., 2007. Nonparametric estimation of a hedonic price function. Journal of Applied Econometrics 22, 695–699.
Phillips, D., 2007. Ethnic and Racial Segregation: A Critical Perspective. Geography Compass 1, 1138–1159.

Racine, J.S., 1997. Consistent significance testing for nonparametric regression. Journal of Business & Economic Statistics 15, 369–378.
Racine, J.S., 2008. Nonparametric econometrics: a primer.
Racine, J.S., Hart, J., Li, Q., 2006. Testing the Significance of Categorical Predictor Variables in Nonparametric Regression Models. Econometric Reviews 25, 523–544.
Racine, J.S., Li, Q., 2004. Nonparametric estimation of regression functions with both categorical and continuous data. Journal of Econometrics 119, 99–130.
Rohe, W.M., Stewart, L.S., 1996. Homeownership and neighborhood stability. Housing Policy Debate 7, 37–81.
Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82, 34.
Rubin, G.M., 1993. Is housing age a commodity? Hedonic price estimates of unit age. Journal of Housing Research 4, 165–184.
Samuelson, P.A., 1938. A note on the pure theory of consumer's behaviour. Economica 61–71.
Samuelson, P.A., 1948. Consumption theory in terms of revealed preference. Economica 243–253.
Schelling, T.C., 1971. Dynamic Models of Segregation. Journal of Mathematical Sociology 1, 143–186.
Schelling, T.C., 1972. The process of residential segregation: Neighborhood tipping, in: Pascal, A.H. (Ed.), Racial Discrimination in Economic Life. Heath, Lexington, MA, pp. 157–184.
Silverman, B.W., 1986. Density estimation for statistics and data analysis. Chapman and Hall, London.
Triplett, J.E., 2006. Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes. OECD Publishing, Paris.
Van Ham, M., Feijten, P., 2008. Who wants to leave the neighbourhood? The effect of being different from the neighbourhood population on wishes to move. Environment and Planning A 40, 1151–1170.
Wilson, W.J., 1987. The truly disadvantaged: The inner city, the underclass, and public policy. University of Chicago Press, Chicago, IL.
Wong, M., 2013. Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore. The Review of Economic Studies.
Yatchew, A., 1998. Nonparametric regression techniques in economics. Journal of Economic Literature 36, 669–721.
Zorlu, A., 2009. Who Leaves the City? The Influence of Ethnic Segregation and Family Ties. Population, Space and Place 15, 323&342.

The UNU-MERIT WORKING Paper Series

2013-01 *Effects of innovation on employment in Latin America* by Gustavo Crespi and Ezequiel Tacsir

2013-02 *Revisiting the porter hypothesis: An empirical analysis of green innovation for the Netherlands* George van Leeuwen and Pierre Mohnen

2013-03 *Impact of external knowledge acquisition strategies on innovation - A comparative study based on Dutch and Swiss panel data* by Spyros Arvanitis, Boris Lokshin, Pierre Mohnen and Martin Wörter

2013-04 *Interactive knowledge exchanges under complex social relations: A simulation modelRobin* by Cowan and Anant Kamath

2013-05 *Innovation systems framework: still useful in the new global context?* by Michiko Iizuka

2013-06 *The importance of intrinsic and extrinsic motivation for measuring IQ* by Lex Borghans, Huub Meijers and Bas ter Weel

2013-07 *Firms' innovation capability-building paths and the nature of changes in learning mechanisms: Multiple case-study evidence from an emerging economy* by Paulo N. Figueiredo , Marcela Cohen and Saulo Gomes

2013-08 *A set of time series data labour market stocks and flows for the Netherlands 1980 to 2010* by Manuel Müllers, Joan Muysken and Erik de Regt

2013-09 *Designing an optimal 'tech fix' path to global climate stability: R&D in a multi-phase climate policy framework* by Adriaan van Zon and Paul A. David

2013-10 *Complementarity between internal knowledge creation and external knowledge sourcing in developing countries* by Jun Hou and Pierre Mohnen

2013-11 *Summarizing large spatial datasets: Spatial principal components and spatial canonical correlation* by Samyukta Bhupathiraju, Bart Verspagen and Thomas Ziesemer

2013-12 *Regional systems of innovation in the Arab region* by Samia Satti Osman Mohamed Nour

2013-13 *Development and social justice: Education, training and health in Sudan* by Samia Satti Osman Mohamed Nour

2013-14 *The economic importance and impacts of intellectual property rights (IPRs) in Sudan* by Samia Satti Osman Mohamed Nour

2013-15 *Overview of knowledge economy in the Arab region* by Samia Satti Osman Mohamed Nour

2013-16 *The importance (impacts) of knowledge at the macro-micro levels in the Arab Gulf countries* by Samia Satti Osman Mohamed Nour

2013-17 *Political determinants and impact analysis of using a cable system as a complement to an urban transport system* by Diego Escobar-García, Francisco García-Orozco and Carlos Cadena-Gaitán

2013-18 *Women entrepreneurs in the informal economy: Is formalization the only solution for business sustainability?* By Shyama V. Ramani, Ajay Thutupalli, Tamas Medovarszki, Sutapa Chattopadhyay, Veena Ravichandran

2013-19 *Heterogeneity in innovation strategies, evolving consumer preferences and market structure: An evolutionary multi-agent based modelling approach* by Salih Çevikarslan

2013-20 *Optimal patent length and patent breadth in an R&D driven market with evolving consumer preferences: An evolutionary multi-agent based modelling approach* by Salih Çevikarslan
2013-21 *Innovation and productivity: An update* by Pierre Mohnen and Bronwyn H. Hall
2013-22 *Fathers' use of parental leave. What do we know?* by Nevena Zhelyazkova
2013-23 *Eliciting Illegal migration rates through list randomization* by David McKenzie and Melissa Siegel
2013-24 *How do ICT firms in Turkey manage innovation? Diversity in expertise versus diversity in markets* by Semih Akçomak, Erdal Akdeve and Derya Fındık
2013-25 *Dynamic models of R&D, innovation and productivity: Panel data evidence for Dutch and French manufacturing* by Wladimir Raymond, Jacques Mairesse, Pierre Mohnen and Franz Palm
2013-26 *Centre-based versus home-based childcare* by Robert Bauchmüller
2013-27 *Microeconometric evidence of financing frictions and innovative activity* by Amaresh K Tiwari, Pierre Mohnen, Franz C Palm and Sybrand Schim van der Loeff
2013-28 *Innovation for economic performance: The case of Latin American firms* by Elena Arias Ortiz, Gustavo Crespi, Ezequiel Tacsir, Fernando Vargas and Pluvia Zuñiga
2013-29 *Is money all? Financing versus knowledge and demand constraints to innovation* by Gabriele Pellegrino and Maria Savona
2013-30 *Child deprivation in Ontario - A (less than perfect) comparison with Europe* by Geranda Notten
2013-31 *Measuring performance: does the assessment depend on the poverty proxy?* by Geranda Notten
2013-32 *How big is the impact of infrastructure on trade? Evidence from meta-analysis* by Mehmet Güney Celbis, Peter Nijkamp and Jacques Poot
2013-33 *Using a 'Systems' Perspective to Explain the Limits of 'New' Multinational Enterprises: the role of 'members-only' location advantages* by Rajneesh Narula
2013-34 *Foreign direct investment as a driver of industrial development: why is there so little evidence?* by Rajneesh Narula
2013-35 *The end of the multifibre arrangement (MFA) and the heterogeneous performance of quota-constrained countries* by Mulu Gebreeyesus
2013-36 *Techological capability building in MNE-related social businesses of less developed countries: The experience of Grameen-Danone Foods in Bangladesh* by Jahan A. Peerally and Paulo N. Figueiredo
2013-37 *The links between economic integration and remittances behaviour of migrants in the Netherlands* by Özge Bilgili
2013-38 *The influence of vulnerability on migration intentions in Afghanistan* by Craig Loschmann and Melissa Siegel
2013-39 *How unemployment insurance savings accounts affect employment duration: Evidence from Chile* by Paula Nagler
2013-40 *Self-organization of knowledge economies* by François Lafond
2013-41 *Designing an optimal 'tech fix' path to global climate stability: Directed R&D and embodied technical change in a multi-phase framework* by Adriaan van Zon & Paul A. David
2013-42 *The growth of outward FDI and the competitiveness of the underlying economy: the case of India* by Rajneesh Narula and Tiju Prasad Kodiyat

2013-43 *The impact of migration on children left behind in Moldova* by Franziska Gassmann, Melissa Siegel, Michaella Vanore and Jennifer Waidler
2013-44 *Technological spillovers and industrial growth in Chinese regions* by Lili Wang, Huub Meijers and Adam Szirmai
2013-45 *Male use of parental leave in Luxembourg: Empirical analysis of administrative records* by Nevena Zhelyazkova
2013-46 *Exploring the paradox of competence-creating subsidiaries: balancing bandwidth and dispersion in MNEs* by Rajneesh Narula
2013-47 *Switching off or switching source: energy consumption and household response to higher energy prices in the Kyrgyz Republic* by Franziska Gassmann and Raquel Tsukada
2013-48 *Beyond technological catch-up: An empirical investigation of further innovative capability accumulation outcomes in latecomer firms with evidence from Brazil* by Paulo N. Figueiredo
2013-49 *Parental leave within the broader work-family trajectory: What can we learn from sequence analysis?* by Nevena Zhelyazkova
2013-50 *Transnational corruption and innovation in transition economies* by Alexis Habiyaremye and Wladimir Raymond
2013-51 *The pace of poverty reduction - A fractional response approach* by Richard Bluhm, Denis de Crombrugghe and Adam Szirmai
2013-52 *Minding Weber more than ever? The impacts of State Capacity and Bureaucratic Autonomy on development goals* by Luciana Cingolani, Kaj Thomsson and Denis de Crombrugghe
2013-53 *The State of State Capacity: a review of concepts, evidence and* measures by Luciana Cingolani
2013-54 *Institutions, Foreign Direct Investment, and Domestic Investment: crowding out or crowding in?* by Kristine Farla, Denis de Crombrugghe and Bart Verspagen
2013-55 *Determinants of firms' investment behaviour: A multilevel approach* by Kristine Farla
2013-56 *Economic development, growth, institutions and geography* by Samyukta Bhupatiraju & Bart Verspagen
2013-57 *Innovation and survival of new firms in Chinese manufacturing, 2000-2006* by Mingqian Zhang and Pierre Mohnen
2013-58 *Government support, innovation and productivity in the Haidian (Beijing) district* by Can Huang, Yilin Wu, Pierre Mohnen and Yanyun Zhao
2013-59 *The impact of the 2009 value added tax reform on enterprise investment and employment - Empirical analysis based on Chinese tax survey data* by Dehua Wang
2013-60 *Doing R&D in a closed or open mode: Dynamics and impacts on productivity* by Julio Miguel Rosa and Pierre Mohnen
2013-61 *Ethnic segregation and heterogeneous preferences of homeowners for housing and neighbourhood characteristics. Evidence from the Netherlands* by Cheng Boon Ong and Kristof De Witte